

Claims

1. Method of automatic processing of a speech signal comprising:

- an automatic step of determination of at least one sequence of probability models coming from a finite directory of models, each sequence describing the probability of acoustic production of a sequence of symbolic units of a phonological nature coming from a finite alphabet, the said sequence of symbolic units corresponding to at least one given text and the said probability models each including an observable random process corresponding to the acoustic production of symbolic units and a non-observable random process having known probability properties, so-called Markov properties;
- a step of determination of a sequence of digital data strings, known as acoustic strings, representing acoustic properties of a speech signal;
- a step of alignment between the said sequence of acoustic strings and the said at least one sequence of models, each model being associated with a sub-sequence of acoustic strings, forming an acoustic segment, and each value of the non-observable process of each model being associated with a sub-sequence of acoustic strings forming an acoustic sub-segment in order to deliver a sequence of non-observable process values associating a value with each acoustic string, known as an aligned sequence; and
- a step of determination of a confidence index of acoustic alignment for each association between a model of the sequence and an acoustic segment, known as a model alignment confidence index, and corresponding to an estimate of the probability *a posteriori* of the model given the observation of the corresponding acoustic segment, known as the *a posteriori* model probability,

characterised in that each step of determination of an alignment confidence index for a model comprises the calculation of the value of the said index at least from a combination of:

- the probability of observation of each acoustic string given the value of the non-observable process, known as the model probability and determined from known characteristic parameters of the probability model;
 - probabilities of production *a priori* of all the models of the said directory, independently of one another, known as the *a priori* model probabilities; and
 - the analytical estimation of the average duration of occupancy of the values of the non-observable process of the model.
2. Method as claimed in Claim 1, characterised in that each step of determination of an acoustic confidence index for a model includes a sub-step of determination of the estimate of the *a priori* probability of each value of the non-observable process of the model, known as the *a priori* value probability, carried out on the basis of the said analytical estimation of the average duration of occupancy of the values of the non-observable process of the model.
3. Method as claimed in Claim 1, characterised in that each step of determination of an alignment confidence index for a model includes a sub-step of determination of a confidence index for each acoustic string forming the acoustic segment associated with the said model and a sub-step of combination of the confidence indices of each string of the said segment in order to deliver the said confidence index of the said model.
4. Method as claimed in Claim 3, characterised in that each sub-step of determination of a confidence index for a given string includes:
- a sub-step of initial calculation combining the model probability, the *a priori* model probability of the model in progress and the average duration of occupancy of the non-observable values for all the values of the non-observable process of the said aligned sequence and of the model in progress;

- a sub-step of calculation of the product of the model probability, the *a priori* model probability and the *a priori* value probability, produced for each value of the non-observable process of all the possible models in the said finite directory of models; and

- a sub-step of summation of all the said products for all the possible models of the said finite directory of models in order to deliver the said confidence index of the said given acoustic string from the results of the said sub-steps.

5. Method as claimed in Claim 1, characterised in that it includes a sub-step of standardisation of the confidence indices by model as a function of the duration of the models.

6. Method as claimed in Claim 1, characterised in that the said automatic step of determination of a sequence of probability models corresponding to a given text includes:

- a sub-step of acquisition of a graphemic representation of the said given text;

- a sub-step of determination of the said sequence of symbolic units from the said graphemic representation; and

- an automatic sub-step of modelling of the said sequence of symbolic units by its breakdown on a base of the said probability models in order to deliver the said sequence of probability models.

7. Method as claimed in Claim 6, characterised in that the said modelling sub-step associates a single probability model with each symbolic unit of the said sequence of symbolic units.

8. Method as claimed in Claim 1, characterised in that the said step of determination of a sequence of digital strings includes:

- a sub-step of acquisition of a speech signal corresponding to the diction of the said given text, adapted in order to deliver a sequence of digital samples of the said speech signal; and

- a sub-step of spectral analysis of the said samples in order to deliver a breakdown of the frequency spectrum of the said speech signal on a non-linear scale, the said breakdown forming the said sequence of acoustic strings.

9. Method as claimed in Claim 8, characterised in that the said sub-step of spectral analysis corresponds to a sub-step of Fourier transformation of the said speech signal, of determination of the distribution of its energy on a non-linear scale by filtering, and of transformation into cosine.

10. Method as claimed in Claim 1, characterised in that the said step of alignment between the said sequence of acoustic strings and the said sequence of models includes:

- a sub-step of calculation of a plurality of possible alignments each associated with a relevance index; and

- a sub-step of selection of a single alignment amongst the said plurality of possible alignments.

11. Method as claimed in Claim 10, characterised in that the said sub-step of determination of a plurality of possible alignments comprises the calculation of at least one optimum alignment, as determined by a so-called Viterbi algorithm.

12. Method as claimed in Claim 1, characterised in that it also includes a step of local modification of the said sequence of models as a function of the said alignment confidence indices determined for each model of the said sequence of models.

13. Method as claimed in Claim 12, characterised in that the said step of local modification comprises a sub-step of deletion of a model from the said sequence of models.

14. Method as claimed in Claim 12, characterised in that the said step of local modification includes a sub-step of substitution of a model of the said sequence of models by another model.

15. Method as claimed in Claim 12, characterised in that the said step of local modification includes a sub-step of insertion of a model between two models of the said sequence of models.

16. Method as claimed in Claim 12, characterised in that the said steps of alignment and of calculation of a confidence index are repeated after each step of local modification of the said sequence of models.

17. Method as claimed in Claim 1, characterised in that the said step of determination of at least one sequence of models is adapted for the determination of a sequence of models corresponding to a given text, and in that the said sequence of acoustic strings represents properties of a speech signal corresponding to the location of the said same given text.

18. Method as claimed in Claim 1, characterised in that the said step of determination of sequences of models is adapted for the determination of a plurality of sequences of models each corresponding to a given text, and in that the said sequence of acoustic strings represents properties of a speech signal corresponding to the location of any text whatsoever, the said method including a step of selection of one or several sequences of models amongst the said plurality for carrying out the said step of determination of confidence indices.

19. Method as claimed in Claim 1, characterised in that the said models are models of which the observable processes have discrete values, the values of the non-observable processes being the states of these processes.

20. Method as claimed in Claim 1, characterised in that the said models are models of which the non-observable processes have continuous values.